An electronic nose in the discrimination of patients with non-small cell lung cancer and COPD

Silvano Dragonieri\textsuperscript{a,b,*}, Jouke T. Annema\textsuperscript{a}, Robert Schot\textsuperscript{a}, Marc P.C. van der Schee\textsuperscript{a,d}, Antonio Spanevello\textsuperscript{c}, Pierluigi Carratu\textsuperscript{b}, Onofrio Resta\textsuperscript{b}, Klaus F. Rabe\textsuperscript{a}, Peter J. Sterk\textsuperscript{a,d}

\textsuperscript{a} Department of Pulmonology, Leiden University Medical Center, Leiden, The Netherlands
\textsuperscript{b} Department of Respiratory Diseases, University of Bari, Bari, Italy
\textsuperscript{c} Department of Respiratory Diseases, Fondazione Maugeri, Cassano Murge, Italy
\textsuperscript{d} Department of Respiratory Medicine, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

Abstract

\textbf{Background:} Exhaled breath contains thousands of gaseous volatile organic compounds (VOCs) that may be used as non-invasive markers of lung disease. The electronic nose analyzes VOCs by composite nano-sensor arrays with learning algorithms. It has been shown that an electronic nose can distinguish the VOCs pattern in exhaled breath of lung cancer patients from healthy controls. We hypothesized that an electronic nose can discriminate patients with lung cancer from COPD patients and healthy controls by analyzing the VOC-profile in exhaled breath.

\textbf{Methods:} 30 subjects participated in a cross-sectional study: 10 patients with non-small cell lung cancer (NSCLC, \( \text{age} 66.4 \pm 9.0, \text{FEV}_1 86.3 \pm 20.7 \)), 10 patients with COPD (\( \text{age} 61.4 \pm 5.5, \text{FEV}_1 70.0 \pm 14.8 \)) and 10 healthy controls (\( \text{age} 58.3 \pm 8.1, \text{FEV}_1 108.9 \pm 14.6 \)). After 5 min tidal breathing through a non-rebreathing valve with inspiratory VOC-filter, subjects performed a single vital capacity maneuver to collect dried exhaled air into a Tedlar bag. The bag was connected to the electronic nose (Cyranose 320) within 10 min, with VOC-filtered room air as baseline. The smellprints were analyzed by onboard statistical software.

\textbf{Results:} Smellprints from NSCLC patients clustered distinctly from those of COPD subjects (cross validation value [CVV]: 85%; M-distance: 3.73). NSCLC patients could also be discriminated from healthy controls in duplicate measurements (CVV: 90% and 80%, respectively; M-distance: 2.96 and 2.26).

\textbf{Conclusion:} VOC-patterns of exhaled breath discriminates patients with lung cancer from COPD patients as well as healthy controls. The electronic nose may qualify as a non-invasive diagnostic tool for lung cancer in the future.

1. Introduction

During the last few years the analysis of exhaled breath has been proposed as a novel option for an early detection of lung cancer. Exhaled breath contains a complex mixture of several hundreds of volatile organic compounds (VOCs) \[1\]. This could be established by gas chromatography \[1\] and mass spectrometry \[2\] (GC–MS). It has been shown that VOCs analysis may be used as a non-invasive marker of lung cancer \[3,4\]. However, the requirement of off-line GC–MS analysis limited the further development of this diagnostic potential.

After the introduction of electronic noses, the sampling of exhaled breath and its VOC-pattern has become readily available, due to their ability to allow on-board analysis and discrimination of “smellprints” by composite nano-sensors arrays (“breathomics”). This is based on pattern recognition without analyzing the individual molecular components \[5\], which potentially suffices for diagnostic objectives.

Recently, the first studies by a sensor array in detecting lung cancer have demonstrated promising diagnostic accuracy \[6–9\]. Therefore, the question arises whether the electronic nose will qualify as a non-invasive diagnostic method for the detection or exclusion of patients with lung cancer.

To date, it is unknown whether concurrent chronic obstructive pulmonary disease (COPD), which is often associated with the development of the neoplasm \[10\], can be responsible for the separate pattern of VOCs in patients with lung cancer. In this study we...
hypothesized that an electronic nose can discriminate the VOCs pattern in exhaled breath between patients with lung cancer and individuals with COPD. Our aim was to address this hypothesis by a cross-sectional study examining the difference in VOC-pattern of exhaled air between patients with a histology-confirmed diagnosis of lung cancer and patients with COPD.

As a secondary aim, we intended to confirm the potential of the electronic nose to distinguish the VOC-patterns between patients with lung cancer and healthy controls and to assess whether this distinction had adequate repeatability.

2. Methods

2.1. Patients

A total number of 30 patients volunteered to participate to this study. All the subjects were adults (45–80 years). The study population included 3 groups of 10 subjects: patients affected by non-small cell lung cancer (NSCLC), patients with chronic obstructive pulmonary disease (COPD) and a healthy control group. Patients were recruited among those visiting the outpatient clinic of the Leiden University Medical Center, whilst controls were recruited by advertisements in the hospital and in the university. Patients with a recent histology established diagnosis of NSCLC and no history of smoking, pre- and post-bronchodilator FEV1/FVC ratio <70% and absence of chronic symptoms or sputum production or dyspnoea at exertion, during the past 4 weeks prior to the measurements and patients with systemic diseases (such as diabetes) or a prior diagnosis of malignancy were not eligible for inclusion in this study. The breathing manoeuvres were standardized, based on a validated method as previously published [15]. In short, patients breathed tidally through a mouthpiece, connected to a 3-way non-rebreathing valve and an inspiratory VOC-filter (A2, North Safety, NL) for 5 min. After a single deep inspiration the patient exhaled a single vital capacity volume into a Tedlar bag connected to the expiratory port. Within 10 min the electronic nose was connected to the Tedlar bag, followed by 1 min sampling of exhaled air through a silica-filled drying chamber [15]. For NSCLC group and controls these manoeuvres were done in duplicate by repeating the same procedure after a 2- to 5-min interval.

2.2. Study design

The study had a cross-sectional case-control design with two visits within a 10-day period. The first day was a screening day to check all the inclusion and exclusion criteria. On the second day exhaled breath was collected in duplicate (except for the COPD group) and sampled by the electronic nose. Smoking was not allowed on the day of the exhaled breath collection and no food and beverages were permitted in the 2 h before the test.

2.3. Electronic nose

We used a Cyranose 320 (Smiths Detections, Pasadena, CA, USA), a handheld portable chemical vapour analyzer, containing a nano-composite array with 32 polymer sensors. When exposed to a gas mixture the sensors swell, thereby changing the electrical resistance, resulting in a unique smellprint of differential electrical resistances [13].

2.4. Lung function

Spirometry (Masterlab Jaeger, Germany) was performed by a trained lung function technician according to the latest recommendations [14] and the forced expiratory volume in 1 s (FEV1) and forced vital capacity (FVC) were measured before and 20 min after 400 μg inhaled salbutamol per metered dose inhaler with a spacer (Volumatic®). Patients withheld short-acting β2 agonists for >8 h and long-acting β2 agonists for >12 h prior to all lung function measurements and electronic nose analysis.

2.5. Exhaled breath collection and sampling

The breathing manoeuvres were standardized, based on a validated method as previously published [15]. In short, patients breathed tidally through a mouthpiece, connected to a 3-way non-rebreathing valve and an inspiratory VOC-filter (A2, North Safety, NL) for 5 min. After a single deep inspiration the patient exhaled a single vital capacity volume into a Tedlar bag connected to the expiratory port. Within 10 min the electronic nose was connected to the Tedlar bag, followed by 1 min sampling of exhaled air through a silica-filled drying chamber [15]. For NSCLC group and controls these manoeuvres were done in duplicate by repeating the same procedure after a 2- to 5-min interval.

2.6. Data analysis

The smellprints were analysed based on recent recommendations [16] by using online software on-board of the electronic nose (PC nose). Data were processed through Savitzky–Golay filtering and analysed by principal component analysis (PCA) and linear canonical discriminant analysis (LDA). A cross validation value (CVV) and the Mahalanobis distance between the group means in units of standard deviation were then calculated. This procedure has been validated in our lab by off-line confirmatory analysis using double cross-validation implementation of linear discriminant analysis on principal component reduction, as previously described [15].

3. Results

The subject characteristics of the three groups are described in Table 1. Patients with lung cancer were slightly older than healthy controls (p < 0.05) whilst no differences in age were reported between the lung cancer group and COPD group. FEV1 (% pred.) was higher in healthy controls as compared to patients with COPD and to those with lung cancer (p < 0.01; p < 0.05, respectively). No significant dif-

| Table 1. Clinical characteristics of the study population |
|-----------------------------------------------|---------|---------|---------|
| Lung cancer | COPD | Healthy controls |
| Patients (n) | 10 | 10 | 10 |
| Age (years) | 66.4 ± 9.0 | 61.4 ± 5.5 | 58.3 ± 8.1 |
| Sex (M/F) | 10/0 | 8/2 | 4/6 |
| FEV1 (% pred) | 86.3 ± 20.7 | 70.0 ± 14.8 | 108.9 ± 14.6 |
| Current smokers (n) | 2 | 6 | 0 |
| Ex-smokers (n) | 7 | 4 | 0 |
| Never smokers (n) | 1 | 0 | 10 |
| Pack-years (n) | n.d. | 30 ± 11 | n.d. |

Values are expressed as mean ± S.D.
ferences in FEV₁ were observed between individuals with lung cancer and COPD.

The clinical characteristics of patients with lung cancer are shown in Table 2. In 6 out of 10 patients the tumour was located within the lung (stage I to II), whereas 4 patients had locally advanced disease (stage IIIA/B NSCLC).

The two-dimensional PCA plot showed that smellprints of patients with lung cancer could be distinguished from those of subjects with COPD (Fig. 1, upper left). Canonical discriminant analysis was then performed on the data (Fig. 1, upper right). With the optimal number of vectors the cross-validation results were 85% correct, with a Mahalanobis distance of 3.73 between the two groups.

Smellprints from lung cancer patients could also be discriminated from those of healthy controls (Fig. 1, lower left). Canonical discriminant analysis gave a cross-validation value of 90% correct and a M-distance of 2.96 between the two groups (Fig. 1, lower right). Analysis of exhaled air from the second bags reproduced these results (lung cancer vs. controls: CVV: 80%; M-distance 2.26).

4. Discussion

Our study demonstrates that an electronic nose can distinguish the pattern of VOCs present in exhaled breath of lung cancer patients from that of patients with COPD. Furthermore the electronic nose could adequately discriminate patients with lung cancer from control subjects. This distinction was confirmed when analysing exhaled air from repeated samples. This indicates that the VOC-patterns in exhaled air differ between two separate smoking-related lung diseases, which warrants further steps towards diagnostic validation of electronic noses in lung cancer and COPD.

The novelty of our study is the comparison of a well-characterized group of NSCLC patients to subjects of similar age with established COPD. To our knowledge, this is the first study that performs a formal group-to-group comparison between patients with lung cancer and COPD. The relevance is given by the fact that this concerns two smoking-related lung diseases and that patients with non-small cell lung cancer are also frequently diagnosed with COPD. Interestingly, we observed a separation of smellprints...

<table>
<thead>
<tr>
<th>Patient number</th>
<th>Histology</th>
<th>TNM</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Squamous cell</td>
<td>T2N1M0</td>
<td>II</td>
</tr>
<tr>
<td>2</td>
<td>NSCLC</td>
<td>T1-2N2M0</td>
<td>IIIA</td>
</tr>
<tr>
<td>3</td>
<td>Adenocarcinoma</td>
<td>T2N0M0</td>
<td>IIIB</td>
</tr>
<tr>
<td>4</td>
<td>Squamous cell</td>
<td>T2N0M0</td>
<td>IIIB</td>
</tr>
<tr>
<td>5</td>
<td>Adenocarcinoma</td>
<td>T1N0M0</td>
<td>I</td>
</tr>
<tr>
<td>6</td>
<td>Adenocarcinoma</td>
<td>T2N3M0</td>
<td>I</td>
</tr>
<tr>
<td>7</td>
<td>NSCLC</td>
<td>T2-4N2M0</td>
<td>IIIA</td>
</tr>
<tr>
<td>8</td>
<td>Squamous cell</td>
<td>T4N2M0</td>
<td>IIIB</td>
</tr>
<tr>
<td>9</td>
<td>Bronchiolo-alveolar T2N0M0</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Squamous cell</td>
<td>T2N0M0</td>
<td>I</td>
</tr>
</tbody>
</table>
between the two groups. Our data extend the previous studies examining lung cancer patients and healthy controls by various sensor array systems [6–9]; Di Natale et al. [6] used quartz microbalance gas sensors showing 100% correct classification of patients with various forms of lung cancer, and 94% classification of controls. Chen et al. [7] presented a study of an electronic nose for detection of lung cancer based on surface acoustic wave sensors and image recognition method. Machado et al. [8] used exactly the same electronic nose as in the present study, and showed adequate discrimination between patients with lung cancer and those from other groups, including healthy controls and a mixed group of subjects with pulmonary diseases (α1-AT deficiency, chronic beryllium disease and COPD). In their validation model these authors found 71.4% sensitivity and 91.9% specificity for detecting bronchogenic carcinoma by an electronic nose, with positive predictive values of 66.6% and 93.4%, respectively. Finally Mazzone et al. [9] used a colorimetric sensor array that predicted the presence of non-small cell lung cancer with a sensitivity of 73.3% and a specificity of 72.4%. Their controls included healthy subjects and patients with COPD, sarcoidosis, pulmonary arterial hypertension and idiopathic pulmonary fibrosis. The present findings suggest that the discrimination between lung cancer and COPD by electronic nose can also be accomplished, which may have pathophysiological and clinical implications.

In our study we carefully considered methodological issues such as the selection of the groups. All the patients and controls were well characterized by internationally standardized and accepted guidelines [11,12]. Such a priori gold-standard diagnosis is required when examining discriminative potential. In addition, we used a validated sampling technique and breathing manoeuvres [15]. The latter is essential, and further optimization of air collection and sampling in the application of electronic noses should be an issue of constant efforts.

Tobacco smoking may play a relevant role in altering the VOC-profile. Previous studies have clearly shown that cigarette smoking directly affects the level of several VOCs in human breath [17]. The smokers in our study abstained from smoking on the day of the test. However, we cannot exclude whether smoking has affected our results. Furthermore, current smoking varied between lung cancer patients and those with COPD and future studies may also have to include groups with similar smoking status. Nevertheless, the comparison of patients with lung cancer and COPD showed a clear distinction between the groups, indicating that we cannot explain our findings only by the presence of smoking-related VOCs or by the presence of smoke-induced airway inflammation.

A potential limitation of our study might be the relatively small sample size. However, 10 subjects per group appeared to be sufficient for obtaining an adequate discrimination between lung cancer patients, COPD patients and controls. In addition, repeated sampling excluded that our results could be explained by accident or by error. This resembles observations in asthma and controls using the same sample sizes [15]. However, the current sample size is not enough to discriminate different stages of lung cancer. It should be noticed that 40% of lung cancer patients in this study had locally advanced disease (stage III). Whether the present findings can be extrapolated to early lung cancer stages needs further investigation.

The implicit limitation of this study is the absence of data about identification of the specific VOCs. Electronic noses are measuring VOC-mixtures in the exhaled breath as a profile, which is analyzed by pattern recognition algorithms [16] to assess whether these smellprints vary between subject groups. Pattern recognition suffices in diagnostic assessment, but specific individual VOC analysis will be essential when examining pathophysiological pathways and the critical individual biomarkers driving the current discrimination. Therefore, the intention of using electronic noses in medicine is not to assess the pathophysiological cause of a disease but to evaluate their diagnostic accuracy for a future application as a clinical test. The identification of specific VOCs by GC–MS provides vital information on the specific molecular pathways involved and is certainly relevant as an aid in optimizing and shaping specific sensors for future clinical purposes.

How can we interpret our findings? It appears that the exhaled breath of subjects with lung cancer is different from that of individuals without it. Several alterations of VOCs have been shown in lung cancer patients using GC–MS [18]. In particular, acetone, butane benzene, decane, isoprene and pentane levels have been found altered in patients with lung cancer [18]. By using solid-phase microextraction combined with gas chromatography, recently Chen et al. [19] distinguished 11 exhaled VOCs as biomarkers of lung cancer, including styrene, decane, isoprene, benzene, undecane, 1-hexene, hexanal, propyl benzene, 1,2,4-trimethyl benzene, heptanal and methyl cyclopentane. Notably, at least two of those (isoprene and undecane) were also the discriminatory VOCs in the headspace of in vitro cultures of lung cancer cells as compared to controls [19]. These metabolites may be driven by increased oxidative events that are associated with the neoplastic processes irrespectively from cigarette smoking [18,20,21]. Moreover, in patients with lung cancer the induction of cytochrome p450 may enhance the catalolysis of VOCs, resulting in an alteration of these compounds [18].

The presence of concurrent COPD may potentially induce a modification in the VOCs-spectrum. COPD is indeed associated with several mechanisms of particular relevance in the development of the disease including proteinase–antiproteinase impairment and chronic airway inflammation sustained by cell recruiting with chemokines secretion and other chemoattractants, leading to oxidative stress and airway remodelling [22]. This hypothesis might be supported by an interesting study by Poli et al. [23], who used mass spectrometry technology to identify lung cancer breath signatures, including a group of patients with COPD as controls. These authors showed that levels of isoprene, 2-methylpentane, ethylbenzene and styrene were statistically different between a group of 36 patients with NSCLC and 25 with COPD. Our results now indicate that despite the presence of potential confounding factors like cigarette smoking and chronic airway inflammation, the electronic nose was able to detect a difference in VOCs pattern in the exhaled breath between patients with lung cancer and those with COPD, strongly suggesting that it might reflect changes in the exhaled breath VOCs-spectrum caused by the non-small cell neoplasm itself or the host response to it.

What is the clinical relevance of our findings? Our data indicate that it is warranted to make the next step in the validation of an electronic nose in diagnostic assessment of lung cancer. This requires specific study designs aimed to test diagnostic accuracy [24]. This includes: (a) the discrimination of a priori diagnosed gold-standard groups (this study), and (b) subsequent external validation in newly recruited, not a priori diagnosed patients with an ‘intention to diagnose’ [24]. If validated in this way, the electronic nose could become a convenient device for the physician because of its non-invasiveness, cheapness and ease to perform. It is self-evident that the definitive diagnosis of lung cancer tissue proof is mandatory. The electronic nose might either be a diagnostic tool for selecting patients for further (invasive) diagnostic procedures or it may qualify as a screening tool in patients with an increased risk of lung cancer to primarily exclude lung cancer. Therefore, the next studies should include a larger population in order to discriminate between patients with various history and stages of lung cancer, and newly presented patients to establish the external validation of future diagnostic accuracy by electronic noses for lung cancer. [24].
Conflict of interest statement

None of the authors has any actual or potential conflict of interest.

Acknowledgements

We thank the patients for volunteering for this study and Stefanie Gauw for helping to recruit the patients.

References